On the Hellmann-Feynman theorem for statistical averages

Francisco M. Fernández

INIFTA (UNLP, CCT La Plata-CONICET), División Química Teórica,
Blvd. 113 y 64 (S/N), Sucursal 4, Casilla de Correo 16,
1900 La Plata, Argentina

Abstract

We discuss the Hellmann-Feynman theorem for degenerate states and its application to the calculation of the derivatives of statistical averages with respect to external parameters.

Some time ago there was a discussion about the validity of the Hellmann-Feynman theorem (HFT)[1] for degenerate states[2–5]. Recently, some of those results[3–5] proved useful in deriving an expression for the derivative of the non-extensive free energy with respect to an external parameter[6]. The author took into account the possible occurrence of degenerate states in the proof of his Lema 1[6]. However, in the proof of his Theorem 2 he appears to assume that the eigenvalues of $\hat{H}$ and of the observable $\hat{A}$ are nondegenerate[6]. We think that this discrepancy should be analyzed carefully. In this letter we investigate the connection between the HFT for quantum-mechanical expectation values and statistical averages when there are degenerate states.

Email address: `fernande@quimica.unlp.edu.ar` (Francisco M. Fernández)

The starting point of our discussion is the Schrödinger equation

$$\hat{H}\psi_m = E_m\psi_m \tag{1}$$

where m is a set of quantum numbers that completely specify the stationary state ψ_m and we assume that $\langle\psi_n|\,\psi_m\rangle = \delta_{mn}$. If the Hamiltonian operator $\hat{H}$ depends on a parameter λ then its eigenvalues and eigenvectors will also depend on it. Following Rastegin[6] we assume that the spectrum of $\hat{H}$ is discrete.

The Hellmann-Feynman theorem for nondegenerate states does not present any difficulty and for this reason we assume that the energy level E_n is g_n-fold degenerate:

$$\hat{H}\psi_{ni} = E_{ni}\psi_{ni},\ E_{ni} = E_n,\ i = 1, 2, \ldots, g_n \tag{2}$$

If we differentiate this equation with respect to λ and then apply the bra $\langle\psi_{nj}|$ from the left, we obtain

$$\langle\psi_{nj}|\,\frac{\partial\hat{H}}{\partial\lambda}\,|\psi_{ni}\rangle = \frac{\partial E_{ni}}{\partial\lambda}\delta_{ij} \tag{3}$$

This equation tells us that there is a set of degenerate eigenvectors for which the diagonal HFT $(i = j)$ is always valid. For simplicity we avoid a detailed discussion of the differentiation of eigenvectors and operators with respect to the external parameter; in this respect we follow earlier approaches to the subject[3–5].

It is convenient to analyse two different cases separately. The simpler one takes place when g_n does not change with λ (at least for all values of physical interest of this external parameter). Any unitary transformation of the degenerate states

$$\chi_i = \sum_{j=1}^{g_n} c_{ji}\psi_{nj},\ i = 1, 2, \ldots, g_n \tag{4}$$

yields a set of g_n eigenvectors of $\hat{H}$ with eigenvalue E_n. They satisfy

$$\langle\chi_i| \frac{\partial \hat{H}}{\partial\lambda} |\chi_j\rangle = \sum_{k=1}^{g_n} c_{ki}^* c_{kj} \frac{\partial E_{nk}}{\partial\lambda} \tag{5}$$

Since g_n does not change with λ it is obvious that $\frac{\partial E_{nk}}{\partial\lambda} = \frac{\partial E_n}{\partial\lambda}$ for all $k = 1, 2, \ldots, g_n$ and this equation simplifies to

$$\langle\chi_i| \frac{\partial \hat{H}}{\partial\lambda} |\chi_j\rangle = \frac{\partial E_{ni}}{\partial\lambda}\delta_{ij} \tag{6}$$

that is similar to (3). In other words: in this case we do not have to worry about choosing a particular set of eigenvectors and all the results derived by Rastegin[6] apply to any observable provided that degeneracy is not removed through variations of λ.

When g_n changes, for example at $\lambda = \lambda_0$, then $\left.\frac{\partial E_{ni}}{\partial\lambda}\right|_{\lambda=\lambda_0} \neq \left.\frac{\partial E_{nj}}{\partial\lambda}\right|_{\lambda=\lambda_0}$ for some $i \neq j$ and Eq. (6) does not follow from Eq. (5). However, in this case we can derive the equation[3]

$$\sum_{i=1}^{g_n} \langle\chi_i| \frac{\partial \hat{H}}{\partial\lambda} |\chi_i\rangle = \sum_{k=1}^{g_n} \frac{\partial E_{nk}}{\partial\lambda} \tag{7}$$

that was invoked by Rastegin[6] to prove his Lemma 1. Typically, $g_n(\lambda) < g_n(\lambda_0)$ which happens, for example, when the symmetry of the system is greater when $\lambda = \lambda_0$.

Before discussing the trace averages that currently appear in statistical mechanics, it is convenient to analyse this problem from another point of view. If we differentiate Eq. (1) with respect to λ and then apply the bra

$\langle\psi_n|$ from the left, we obtain an expression for both the diagonal ($m = n$) and off-diagonal ($m \neq n$) HFT[7]

$$\langle\psi_n| \frac{\partial \hat{H}}{\partial \lambda} |\psi_m\rangle = (E_m - E_n) \langle\psi_n| \left.\frac{\partial \psi_m}{\partial \lambda}\right\rangle + \frac{\partial E_m}{\partial \lambda}\delta_{mn} \tag{8}$$

If the eigenvalues E_m and E_n are degenerate at $\lambda = \lambda_0$ $E_m(\lambda_0) = E_n(\lambda_0)$ then

$$\left.\langle\psi_n| \frac{\partial \hat{H}}{\partial \lambda} |\psi_m\rangle\right|_{\lambda=\lambda_0} = \left.\frac{\partial E_m}{\partial \lambda}\right|_{\lambda=\lambda_0} \delta_{mn} \tag{9}$$

This equation is identical to Eq. (3) but its derivation reveals that the diagonal HFT applies to degenerate states provided that we choose the eigenvectors of $\hat{H}$ according to

$$\psi_n(\lambda_0) = \lim_{\lambda\to\lambda_0} \psi_n(\lambda) \tag{10}$$

In what follows we analyse the HFT in the context of statistical-averages that we develop in a somewhat more general setting than that considered by Rastegin[6]. For any Hermitian operator $\hat{W}$ that commutes with $\hat{H}$:

$$\left[\hat{H}, \hat{W}\right] = 0 \tag{11}$$

the hypervirial theorem

$$\langle\psi_i| \left[\hat{H}, \hat{W}\right] |\psi_j\rangle = (E_i - E_j) \langle\psi_i| \hat{W} |\psi_j\rangle \tag{12}$$

tells us that

$$\langle\psi_i| \hat{W} |\psi_j\rangle = 0 \;\; \text{if } E_i \neq E_j \tag{13}$$

If the trace

$$tr\left(\hat{W}\frac{\partial \hat{H}}{\partial \lambda}\right) = \sum_i \sum_j \langle\varphi_i| \hat{W} |\varphi_j\rangle \langle\varphi_j| \frac{\partial \hat{H}}{\partial \lambda} |\varphi_i\rangle \tag{14}$$

exists then it is invariant under unitary transformations of the basis set and we can thus choose the eigenvectors of $\hat{H}$ (10) that satisfy Eq. (9). Since they also satisfy Eq. (13) we have

$$tr\left(\hat{W}\frac{\partial\hat{H}}{\partial\lambda}\right)=\sum_n\left\langle\psi_n\right|\hat{W}\left|\psi_n\right\rangle\left\langle\psi_n\right|\frac{\partial\hat{H}}{\partial\lambda}\left|\psi_n\right\rangle=\sum_n\left\langle\psi_n\right|\hat{W}\left|\psi_n\right\rangle\frac{\partial E_n}{\partial\lambda} \tag{15}$$

This expression is the basis for many of the results derived by Rastegin[6] such as, for example, his Lemma 1:

$$tr\left[f(\hat{H})\frac{\partial\hat{H}}{\partial\lambda}\right]=\sum_n f(E_n)\frac{\partial E_n}{\partial\lambda} \tag{16}$$

We appreciate that we do not have to worry about degeneracy when calculating traces provided that $\hat{W}$ is diagonal with respect to the nondegenerate eigenvectors of $\hat{H}$. In other words: we do not need to invoke the HFT sum expression (7).

Equation (15) also applies to any operator $\hat{A}$ that depends on a parameter λ, exhibits a discrete spectrum and commutes with $\hat{W}$. Following Rastegin[6] we choose an element of the complete set of commuting observables that shares a common eigenbasis with $\hat{H}$

$$\hat{A}\psi_m=a_m\psi_m \tag{17}$$

In such a case we have

$$tr\left(\hat{W}\frac{\partial\hat{A}}{\partial\lambda}\right)=\sum_n\left\langle\psi_n\right|\hat{W}\left|\psi_n\right\rangle\frac{\partial a_n}{\partial\lambda} \tag{18}$$

provided that

$$\left.\left\langle\psi_n\right|\frac{\partial\hat{A}}{\partial\lambda}\left|\psi_m\right\rangle\right|_{\lambda=\lambda_0}=\left.\frac{\partial a_m}{\partial\lambda}\right|_{\lambda=\lambda_0}\delta_{mn} \tag{19}$$

when $E_m(\lambda_0)=E_n(\lambda_0)$ and $E_m(\lambda)\neq E_n(\lambda)$ for $\lambda\neq\lambda_0$.

Rastegin's equations (27) and (31) that are necessary for proving his theorem 2[6] require that the eigenvectors satisfy present equations (9) and (19) when g_n changes at λ_0. If one does not state these conditions explicitly then one is in principle assuming that g_n does not change with λ and the resulting theorems are not so widely applicable.

References

[1] R. P. Feynman, Forces in Molecules, Phys. Rev. 56 (1939) 340-343.

[2] G. P. Zhang and T. F. George, Breakdown of the Hellmann-Feynman theorem: Degeneracy is the key, Phys. Rev. B 66 (2002) 033110.

[3] F. M. Fernández, Comment on "Breakdown of the Hellmann-Feynman theorem: Degeneracy is the key", Phys. Rev. B 69 (2004) 037101.

[4] S. R. Vatsya, Comment on "Breakdown of the Hellmann-Feynman theorem: Degeneracy is the key", Phys. Rev. B 69 (2004) 037102.

[5] R. Balawender, A. Holas, and N. H. March, Comment on "Breakdown of the Hellmann-Feynman theorem: Degeneracy is the key", Phys. Rev. B 69 (2004) 037103.

[6] A. E. Rastegin, Formulation of the Hellmann-Feynman theorem for the "second choice" version of Tsallis' thermostatistics, Physica A 392 (2013) 103-110.

[7] J. C. Y. Chen, General relation of fundamental molecular constants between two electronic states, J. Chem. Phys. 38 (1963) 832-839.